

# Convergence of Sparse Collocation for Functions of Countably Many Gaussian Random Variables (with Application to Lognormal Elliptic Diffusion Problems)

Oliver G. Ernst<sup>†</sup>, Björn Sprungk<sup>†</sup> and Lorenzo Tamellini<sup>\*</sup>

<sup>†</sup> Department of Mathematics, TU Chemnitz, Germany

<sup>\*</sup> Istituto di Matematica Applicata e Tecnologie Informatiche “E. Magenes” del CNR, Pavia, Italy

## Abstract

We give a convergence proof of sparse collocation to approximate Hilbert space-valued functions depending on countably many Gaussian random variables. Such functions appear as solutions or quantities of interest associated with elliptic PDEs with lognormal diffusion coefficients. We outline a general  $L^2$ -convergence theory based on previous work by Bachmayr et al. [4] and Chen [9] and establish an algebraic convergence rate for sufficiently smooth functions assuming a mild growth bound for the univariate hierarchical surpluses of the interpolation scheme applied to Hermite polynomials. We verify specifically for Gauss-Hermite nodes that this assumption holds and also show algebraic convergence w.r.t. the resulting number of sparse grid points for this case.

**Mathematics Subject Classification:** 65D05, 65D15, 65C30, 60H25.

## 1 Introduction

The elliptic diffusion problem

$$-\nabla \cdot (a(\omega) \nabla u(\omega)) = f \quad \text{in } D \subset \mathbb{R}^d, \quad u(\omega) = 0 \text{ on } \partial D \quad (1)$$

with a random diffusion coefficient  $a : \Omega \rightarrow L^\infty(D)$  has become the standard model problem for numerical methods for solving random PDEs. For modelling reasons the diffusion random field is often taken to have a lognormal probability law, which complicates both the study of the well-posedness of the problem [8, 22, 17, 28] as well as the analysis of approximation methods. One of the challenges is that the most common parametrization of a Gaussian random field – the *Karhunen-Loève expansion* [2, 20] – involves a countable number of standard normal random variables

$$\log a(x, \omega) = \phi_0(x) + \sum_{m=1}^{\infty} \phi_m(x) \xi_m(\omega), \quad \phi_m \in L^\infty(D), \quad \xi_m \sim N(0, 1) \text{ i.i.d.}, \quad m \in \mathbb{N}, \quad (2)$$

leading to an elliptic PDE with a countably infinite number of random parameters  $\xi = (\xi_m)_{m \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ .

Besides the stochastic Galerkin method [20, 27] the most common methods for approximating the solution  $u(\xi)$  of such random or parametric elliptic PDEs are probably *polynomial collocation methods*. Early works on such methods for random PDEs considered a finite, if large, number of random parameters, a setting also referred to as *finite-dimensional noise* [38, 3, 34, 33]. In this case the parametric representation of  $\log a$  is typically obtained by truncating a series expansion of the random field such as (2).

The analysis of the problem depending on an infinite number of random variables was first discussed by Cohen, DeVore and Schwab in [14, 15], in the simpler setting in which the diffusion field  $a$  itself is expanded in series (instead of its log as in (2)), thus obtaining an affine dependence of  $a$  over the random variables  $\xi_m$ , which are moreover assumed to have bounded support. In this framework, the convergence of the best  $N$ -term approximation of the solution of the diffusion equation by Taylor and Legendre series was shown to be independent of the number of random variables; this result was further refined in the recent paper [5]. Employing the theoretical concepts stated in [14, 15] Chkifa, Cohen and Schwab analyze in [11] collocation methods based on Lagrange interpolation with Leja points for problems with diffusion coefficient depending linearly on an infinite number of bounded random variables, which are adaptive in the polynomial degree as well as the number of active dimensions or random variables, respectively. The adaptive algorithm itself is related to the earlier work [19]. Each interpolatory approximation gives rise to a quadrature scheme, and in [35] Schillings and Schwab consider sparse adaptive quadrature schemes in the same setting of [11] in connection with approximating expectations with respect to posterior measures in Bayesian inference. Extensions to the case where the diffusion coefficient  $a$  depends non-linearly on an infinite number of random variables with bounded support was instead discussed in [12].

Returning to the original lognormal diffusion problem, i.e. with  $a$  being expanded as in equation (2) and depending on random variables with unbounded support, Hoang and Schwab [24] have obtained convergence results on best  $N$ -term approximations via Hermite polynomials. These were recently extended by Bachmayr et al. [4] by a different analytical approach employing a weighted  $\ell^2$ -summability of the coefficients of the Hermite expansion of the solution and their relation to partial derivatives. The theoretical tools provided in [4] enabled then a convergence analysis for adaptive sparse quadrature [9] employing, e.g., Gauss-Hermite nodes for Banach space-valued functions of countably many Gaussian random variables. Hence, it remains to analyze the convergence of sparse polynomial collocation for functions of infinitely many Gaussian random variables, such as the solution to the lognormal diffusion problem (1).

In this paper, we follow the approach of [4] and [9] to prove an algebraic convergence rate with respect to the number of nodes for adaptive sparse grid collocation based on Gauss-Hermite nodes in the case of countably many variables. In particular, the result applies to solutions  $u$  of (1) with  $a$  a lognormal random field. In addition, we highlight the common ideas surrounding sparse collocation found in the works mentioned above. In particular, note that the convergence result in terms of the number of nodes is obtained in two steps, i.e. first linking the error to the size of the multiindex set defining the sparse grid, and then expressing a bound on the number of points of such grid: this procedure has been followed also in all the above-mentioned works analyzing the convergence of sparse grid quadrature and collocation schemes. An alternative strategy, that links instead directly the error to the number of collocation points by introducing the so-called “profits” of each component of the sparse grids, has been discussed in [31, 23], albeit only in the case of random variables with bounded support.

We remark that besides the classical node families such of Gauss-Hermite and Genz-Keister [18] for quadrature and interpolation on  $\mathbb{R}$  with respect to a Gaussian weight, Jakeman and Narayan [29] have introduced the concept of *weighted Leja points* – a generalization of the classical Leja points (see e.g. [16, 10] and references therein) to unbounded domains and arbitrary weight functions. Moreover, they have proved that these node sets possess the correct asymptotic distribution of interpolation nodes and illustrate their computational potential in numerical experiments. Note that such weighted Leja points provide a nested and linearly growing sequence of interpolation nodes. The analysis of sparse grid collocation based on normal Leja points, i.e., weighted Leja points for a Gaussian weight, is an interesting topic for future research.

The remainder of the paper is organized as follows. In the next section we introduce the general setting and notation and construct the sparse grid collocation operator based on univariate Lagrange interpolation operators. Section 3 is devoted to the convergence analysis of such operators. First, we outline in Subsection 3.1 the general approaches to prove algebraic convergence rates as they can be found in the works mentioned above. Later, we follow in Subsection 3.2 the approach of [4, 9] and derive sufficient conditions for the un-

derlying univariate interpolation nodes in order to obtain such rates when approximating “countably-variate” functions of certain smoothness. Finally, in Subsection 3.3 we verify these conditions for Gauss-Hermite nodes, provide sharp estimates for the number of nodes in the resulting sparse grids, and state a convergence result with respect to this number. Section 4 comes back to our motivation and comments on the application to random elliptic PDEs before we draw final conclusions.

## 2 Setting and Sparse Grid Collocation

We consider functions  $f$  defined on a parameter domain  $\Gamma \subseteq \mathbb{R}^N$  taking values in a separable real Hilbert space  $\mathcal{H}$  with inner product  $(\cdot, \cdot)_{\mathcal{H}}$  and norm  $\|\cdot\|_{\mathcal{H}}$ . As our interest lies in the approximation of the dependence of  $f : \Gamma \rightarrow \mathcal{H}$  on  $\xi \in \Gamma$  by multivariate polynomials based on Lagrange interpolation, a minimal requirement is that point evaluation of  $f$  at any  $\xi \in \Gamma$  be well-defined. Stronger smoothness requirements on  $f$  become necessary when deriving convergence rate estimates for the approximations.

We introduce a probability measure  $\mu$  on  $\mathbb{R}^N$  as the countable product measure of standard Gaussian measures on  $\mathbb{R}$ , i.e.,

$$\mu = \bigotimes_{m \geq 1} N(0, 1). \quad (3)$$

and denote by  $L_{\mu}^2(\Gamma; \mathcal{H})$  the space of all (equivalence classes of) functions with finite second moments with respect to  $\mu$  in the sense that

$$\int_{\mathbb{R}^N} \|f(\xi)\|_{\mathcal{H}}^2 \mu(d\xi) < \infty$$

which forms a Hilbert space with inner product

$$(f, g) = \int_{\mathbb{R}^N} (f(\xi), g(\xi))_{\mathcal{H}} \mu(d\xi).$$

In the following we require

**Assumption A1.** Let  $\Gamma \subset \mathbb{R}^N$  be an open domain with  $\mu(\Gamma) = 1$  and let  $f : \Gamma \rightarrow \mathcal{H}$ . There holds (for a measurable extension of  $f$  to  $\mathbb{R}^N$ ) that  $f \in L_{\mu}^2(\mathbb{R}^N; \mathcal{H})$ .

It is shown e.g. in [36, Theorem 2.5], that the countable tensor product of Hermite polynomials forms an orthonormal basis of  $L_{\mu}^2(\mathbb{R}^N; \mathcal{H})$ . Under Assumption A1 we therefore have

$$f(\xi) = \sum_{\nu \in \mathcal{F}} f_{\nu} H_{\nu}(\xi), \quad f_{\nu} := \int_{\mathbb{R}^N} f(\xi) H_{\nu}(\xi) \mu(d\xi) \in \mathcal{H}, \quad (4)$$

where  $H_{\nu}(\xi) = \prod_{m \geq 1} H_{\nu_m}(\xi_m)$  and  $H_{\nu}$  denotes the univariate Hermite orthonormal polynomial of degree  $\nu$  as well as

$$\mathcal{F} := \left\{ \nu \in \mathbb{N}_0^N : |\nu|_0 < \infty \right\}, \quad |\nu|_0 := |\{j \in \mathbb{N} : \nu_j > 0\}|.$$

### 2.1 Sparse Polynomial Collocation

The construction of sparse collocation operators below is based on sequences of univariate Lagrange interpolation operators  $U_k$  mapping into the set  $\mathcal{P}_k$  of univariate polynomials of degree at most  $k \in \mathbb{N}_0$ . Thus,

$$(U_k f)(\xi) = \sum_{i=0}^k f(\xi_i^{(k)}) L_i^{(k)}(\xi), \quad f : \mathbb{R} \rightarrow \mathbb{R},$$

where  $\{L_i^{(k)}\}_{i=0}^k$  denote the Lagrange fundamental polynomials of degree  $k$  associated with the set of  $k+1$  distinct interpolation nodes  $\Xi^{(k)} = \{\xi_0^{(k)}, \xi_1^{(k)}, \dots, \xi_k^{(k)}\}$ .

**Remark 1.** It may also be of interest to consider sequences of interpolation operators  $U_i$  with a more general nondecreasing polynomial exactness degree  $n(k)$  with  $n : \mathbb{N}_0 \rightarrow \mathbb{N}_0$  and  $n(0) = 0$ . However, we restrict ourselves to  $n(k) = k$  for simplicity.

Introducing the *detail operators*

$$\Delta_k := U_k - U_{k-1}, \quad k \geq 0,$$

where we set  $U_{-1} := 0$ , we observe that

$$U_k = U_{k-1} + \Delta_k = \Delta_0 + \Delta_1 + \cdots + \Delta_k.$$

**Tensorization.** For any multi-index  $\mathbf{k} = (k_m)_{m \in \mathbb{N}} \in \mathcal{F}$  the (full) tensor product interpolation operator  $U_{\mathbf{k}} := \bigotimes_{m \in \mathbb{N}} U_{k_m}$  is defined by

$$(U_{\mathbf{k}} f)(\boldsymbol{\xi}) = \left[ \bigotimes_{m \in \mathbb{N}} U_{k_m} f \right](\boldsymbol{\xi}) = \sum_{\mathbf{i} \leq \mathbf{k}} f(\boldsymbol{\xi}_i^{(\mathbf{k})}) L_i^{(\mathbf{k})}(\boldsymbol{\xi}), \quad f : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}, \quad (5)$$

where  $\boldsymbol{\xi}_i^{(\mathbf{k})} \in \mathbb{R}^{\mathbb{N}}$  ranges over all points in the Cartesian product

$$\Xi^{(\mathbf{k})} := \bigtimes_{m \in \mathbb{N}} \Xi^{(k_m)}, \quad \text{with} \quad |\Xi^{(\mathbf{k})}| = \prod_{m \in \mathbb{N}} (1 + k_m), \quad (6)$$

and

$$L_i^{(\mathbf{k})}(\boldsymbol{\xi}) := \prod_{m \in \mathbb{N}} L_{i_m}^{(k_m)}(\xi_m)$$

is a multivariate polynomial of (total) degree  $|\mathbf{k}|_1 = \sum_m k_m$ . Note that  $L_i^{(\mathbf{k})}(\boldsymbol{\xi}) \equiv 1$  if  $\mathbf{k} = 0$ ; in particular, since  $\mathbf{k} \in \mathcal{F}$  only a finite numbers of factors are  $\neq 1$  so that these products can be regarded as finite. We denote this multivariate (tensor-product) polynomial space by

$$\mathcal{Q}_{\mathbf{k}} := \text{span}\{\boldsymbol{\xi}^i : 0 \leq i_m \leq k_m, m \in \mathbb{N}\}, \quad \mathbf{k} \in \mathcal{F}. \quad (7)$$

Note that, since both the univariate polynomial sets of Lagrange fundamental polynomials  $\{L_i^{(k)}\}_{i=0}^k$  and the Hermite orthonormal polynomials  $\{H_i\}_{i=0}^k$  form a basis of  $\mathcal{P}_k$ , equivalent characterizations are

$$\begin{aligned} \mathcal{Q}_{\mathbf{k}} &= \text{span}\{L_i^{(\mathbf{k})} : 0 \leq i_m \leq k_m, m \in \mathbb{N}\} \\ &= \text{span}\{H_i : 0 \leq i_m \leq k_m, m \in \mathbb{N}\}, \quad \mathbf{k} \in \mathcal{F}. \end{aligned}$$

In order for the tensor product interpolation operator  $U_{\mathbf{k}}$  to be applicable also to functions defined only on a subset  $\Gamma \subset \mathbb{R}^{\mathbb{N}}$ , we assume the interpolation nodes to all lie in  $\Gamma$ :

**Assumption A2.** Let  $\Gamma \subset \mathbb{R}^{\mathbb{N}}$  denote the open domain from Assumption A1. For all  $\mathbf{k} \in \mathcal{F}$  the Cartesian products of nodal sets  $\Xi^{(\mathbf{k})}$  given in (6) satisfy  $\Xi^{(\mathbf{k})} \subset \Gamma$ .

In the following we denote by  $\mathbb{R}^{\Gamma}$  the set of all mappings from  $\Gamma$  to  $\mathbb{R}$ . In analogy to (5) we define for any multi-index  $\mathbf{k} \in \mathcal{F}$  the tensorized detail operator

$$\Delta_{\mathbf{k}} := \bigotimes_{m \in \mathbb{N}} \Delta_{k_m} : \mathbb{R}^{\Gamma} \rightarrow \mathcal{Q}_{\mathbf{k}}.$$

Finally, we associate with a finite subset  $\Lambda \subset \mathcal{F}$  the multivariate polynomial space

$$\mathcal{P}_\Lambda := \text{span}\{\xi^\nu : \nu \in \Lambda\} \quad (8)$$

as well as the approximation operator  $U_\Lambda : \mathbb{R}^\Gamma \rightarrow \mathcal{P}_\Lambda$  by

$$U_\Lambda := \sum_{i \in \Lambda} \Delta_i. \quad (9)$$

We will see that  $U_\Lambda$  is exact on  $\mathcal{P}_\Lambda$  under some natural assumptions on the multi-index set  $\Lambda$ , for which we first recall some basic definitions given in [13, 11, 12].

**Partial orderings and monotone sets of multiindices.** We define a partial ordering on  $\mathcal{F}$  by

$$\tilde{\nu} \leq \nu \quad :\Leftrightarrow \quad \tilde{\nu}_m \leq \nu_m \quad \forall m \in \mathbb{N}$$

as well as

$$\tilde{\nu} < \nu \quad :\Leftrightarrow \quad \tilde{\nu} \leq \nu \text{ and } \tilde{\nu}_m < \nu_m \text{ for at least one } m \in \mathbb{N}$$

and introduce the relation

$$\tilde{\nu} \not\leq \nu \quad :\Leftrightarrow \quad \tilde{\nu}_m > \nu_m \text{ for at least one } m \in \mathbb{N}.$$

We shall call a set of multi-indices  $\Lambda \subset \mathcal{F}$  *monotone* iff  $\tilde{\nu} \leq \nu$  for any  $\tilde{\nu} \in \mathcal{F}$  and  $\nu \in \Lambda$  implies that also  $\tilde{\nu} \in \Lambda$ . Finally, for a multiindex  $\nu \in \mathcal{F}$  we define its *rectangular envelope*  $\mathcal{R}_\nu$  by

$$\mathcal{R}_\nu := \{\tilde{\nu} \in \mathcal{F} : \tilde{\nu} \leq \nu\}.$$

Note that  $\mathcal{R}_\nu$  for  $\nu \in \mathcal{F}$  is a finite set with cardinality

$$|\mathcal{R}_\nu| = \prod_{m \in \mathbb{N}} (1 + \nu_m) < \infty. \quad (10)$$

The introduction of the rectangular envelope  $\mathcal{R}_\nu$  of a multi-index  $\nu \in \mathcal{F}$  permits a convenient characterization of the polynomial spaces  $\mathcal{Q}_k$  and  $\mathcal{P}_\Lambda$  introduced in (7) and (8).

**Proposition 2.** If  $\Lambda \subset \mathcal{F}$  is monotone, then

$$\Lambda = \bigcup_{\nu \in \Lambda} \mathcal{R}_\nu.$$

*Proof.* Since  $\nu \in \mathcal{R}_\nu$  for all  $\nu \in \Lambda$  the set on the left is obviously a subset of that on the right. Conversely, given  $i \in \mathcal{R}_\nu$  for some  $\nu \in \Lambda$ , the definition of  $\mathcal{R}_\nu$  implies  $i \leq \nu$ , which in turn implies  $i \in \Lambda$  by the monotonicity of  $\Lambda$ .  $\square$

In particular, we conclude

$$\begin{aligned} \mathcal{P}_\Lambda &= \text{span}\{\xi^i : i \in \mathcal{R}_\nu, \nu \in \Lambda\} = \sum_{\nu \in \Lambda} \mathcal{Q}_\nu \\ &= \text{span}\{H_\nu : \nu \in \Lambda\} = \text{span}\{L_i^{(\nu)} : i \leq \nu, \nu \in \Lambda\}, \end{aligned}$$

and in this sense  $\mathcal{P}_\Lambda$  for a multi-index set  $\Lambda \subset \mathcal{R}_k$  represents a sparsification of  $\mathcal{Q}_k$ . In particular, the full tensor product polynomial space  $\mathcal{Q}_k$  coincides with  $\mathcal{P}_\Lambda$  for  $\Lambda = \mathcal{R}_k$ . Similarly, the full tensor approximation operator  $U_k$  defined in (5) can be expressed as  $U_k = \sum_{i \in \mathcal{R}_k} \Delta_i$ .

**Proposition 3.** Let  $\Lambda \subset \mathcal{F}$  be a finite and monotone set. Then  $U_\Lambda p = p$  for all  $p \in \mathcal{P}_\Lambda$ . In particular, for all  $p \in \mathcal{P}_\Lambda$  we have  $\Delta_i p = 0$  for  $i \notin \Lambda$ .

*Proof.* Observe first that, for any  $\nu, i \in \mathcal{F}$  such that  $i \not\leq \nu$  we have

$$\Delta_i \xi^\nu = \prod_{m \in \mathbb{N}} \Delta_{i_m} \xi_m^{\nu_m} = \prod_{m \in \mathbb{N}} \underbrace{(U_{i_m} - U_{i_m-1}) \xi_m^{\nu_m}}_{= \xi_m^{\nu_m} - \xi_m^{\nu_m-1} \equiv 0 \text{ for at least one } m} = 0.$$

It suffices to prove the assertions for all monomials  $\xi^\nu$  in  $\mathcal{P}_\Lambda$ . For  $\nu \in \Lambda$  any  $i \in \mathcal{F} \setminus \Lambda$  must satisfy  $i \not\leq \nu$  and therefore  $\Delta_i \xi^\nu = 0$ , proving the second assertion. We conclude that

$$U_\Lambda \xi^\nu = \sum_{i \in \Lambda} \Delta_i \xi^\nu = \sum_{i \in \Lambda \cap \mathcal{R}_\nu} \Delta_i \xi^\nu = \sum_{i \in \mathcal{R}_\nu} \Delta_i \xi^\nu,$$

where the third equality follows from the fact that  $\mathcal{R}_\nu \subseteq \Lambda$  for all  $\nu \in \Lambda$  due to the monotonicity of  $\Lambda$ . The proof concludes with

$$U_\Lambda \xi^\nu = \sum_{i \in \mathcal{R}_\nu} \Delta_i \xi^\nu = \sum_{i \in \mathcal{R}_\nu} \left( \prod_{m \in \mathbb{N}} \Delta_{i_m} \xi_m^{\nu_m} \right) = \prod_{m \in \mathbb{N}} \left( \sum_{i_m=0}^{\nu_m} \Delta_{i_m} \xi_m^{\nu_m} \right) = \prod_{m \in \mathbb{N}} U_{\nu_m} \xi_m^{\nu_m} = \prod_{m \in \mathbb{N}} \xi_m^{\nu_m} = \xi^\nu.$$

Note that the third equality is obtained by rewriting a (finite) product of sums: since  $\nu \in \mathcal{F}$  there exists an  $M \in \mathbb{N}$  such that  $\nu_m = 0$  for  $m > M$ . For such  $m$  we have  $\Delta_{i_m} \xi_m^{\nu_m} = \Delta_0 \xi_m^0 \equiv 1$  and therefore

$$\begin{aligned} \prod_{m \in \mathbb{N}} \left( \sum_{i_m=0}^{\nu_m} \Delta_{i_m} \xi_m^{\nu_m} \right) &= (\Delta_0 \xi_1^{\nu_1} + \dots + \Delta_{\nu_1} \xi_1^{\nu_1}) \dots (\Delta_0 \xi_M^{\nu_M} + \dots + \Delta_{\nu_M} \xi_M^{\nu_M}) \\ &= \sum_{\substack{i \in \mathbb{N}_0^M \\ i_m \leq \nu_m}} \Delta_{i_1} \xi_1^{\nu_1} \dots \Delta_{i_M} \xi_M^{\nu_M} = \sum_{i \in \mathcal{R}_\nu} \left( \prod_{m \in \mathbb{N}} \Delta_{i_m} \xi_m^{\nu_m} \right). \end{aligned}$$

□

Proposition 3 can be seen as an extension of [6, Proposition 1] to general monotone multi-index sets as well as an extension of [13, Theorem 6.1] and [11, Theorem 2.1] to interpolation operators  $U_i$  with non-nested node sets. As mentioned in [13, p. 89], if the set  $\Lambda$  is not monotone then  $U_\Lambda$  will not be exact on  $\mathcal{P}_\Lambda$  in general. However, the exactness on  $\mathcal{P}_\Lambda$  is a crucial property in the subsequent convergence analysis and we therefore choose to work exclusively with monotone sets  $\Lambda$ . The unisolvence on  $\mathcal{P}_\Lambda$  of point evaluation on the multivariate node set  $\Xi_\Lambda := \{\xi_\nu : \nu \in \Lambda\}$  is discussed in [13, Theorem 6.1].

### 3 Convergence Analysis

In this section we analyse the error

$$\|f - U_\Lambda f\|_{L_\mu^2}, \quad f: \Gamma \rightarrow \mathcal{H},$$

where  $\|\cdot\|_{L_\mu^2}$  denotes the norm in  $L_\mu^2(\mathbb{R}^{\mathbb{N}}; \mathcal{H})$ ,  $f$  is assumed to satisfy Assumption A1 and  $\Lambda \subset \mathcal{F}$  is required to be monotone. Our main goal here is to establish a convergence rate  $s > 0$  for the error of  $U_{\Lambda_N}$  for a nested sequence  $\Lambda_N$  of monotone subsets of  $\mathcal{F}$  with  $|\Lambda_N| = N$ , i.e.,

$$\|f - U_{\Lambda_N} f\|_{L_\mu^2} \leq CN^{-s}, \quad f: \Gamma \rightarrow \mathcal{H}, \quad (11)$$

where  $C < \infty$  may depend on  $f$  as well as the univariate nodal sets. The line of proof we present here follows and builds upon the works [9, 24, 4].

### 3.1 General Convergence Results

The subsequent error analysis for the sparse collocation operator  $U_\Lambda$  is based on the representation of multivariate functions  $f \in L_\mu^2(\mathbb{R}^N; \mathcal{H})$  in the orthonormal basis of multivariate Hermite polynomials  $H_\nu$ . We shall therefore examine the worst-case approximation error of any  $U_\Lambda$  applied to a given multivariate Hermite basis polynomial  $H_\nu$ . To this end we define

$$c_\nu := \sup_{\Lambda \subset \mathcal{F}, |\Lambda| < \infty} \|(I - U_\Lambda)H_\nu\|_{L_\mu^2}, \quad \nu \in \mathcal{F}. \quad (12)$$

This quantity is finite since  $\Delta_i H_\nu = 0$  for  $i \not\leq \nu$  and hence

$$\sup_{\Lambda \subset \mathcal{F}, |\Lambda| < \infty} \|(I - U_\Lambda)H_\nu\|_{L_\mu^2} = \max_{\Lambda \subseteq \mathcal{R}_\nu} \|(I - U_\Lambda)H_\nu\|_{L_\mu^2},$$

where the maximum is taken over a finite set. The quantities  $c_\nu$  also measure the deviation of the error of oblique projection  $U_\Lambda$  from that of orthogonal projection, as these numbers would all be zero or one if  $U_\Lambda$  is replaced with the  $L_\mu^2$ -orthogonal projection onto  $\mathcal{P}_\Lambda$ . Moreover, we obtain the following bound:

**Proposition 4.** For all  $\nu \in \mathcal{F}$  the quantity  $c_\nu$  defined in (12) satisfies

$$c_\nu \leq \sum_{i \in \mathcal{R}_\nu} \|\Delta_i H_\nu\|_{L_\mu^2}.$$

In particular, if there holds for a  $\theta > 0$  and a  $K \geq 1$ ,

$$\|\Delta_i H_\nu\|_{L_\mu^2} \leq (1 + K\nu)^\theta \quad \text{for all } i \in \mathbb{N}_0, \quad (13)$$

where we have denoted the univariate Gaussian measure again by  $\mu$ , then

$$c_\nu \leq \prod_{m \in \mathbb{N}} (1 + K\nu_m)^{\theta+1}, \quad \nu \in \mathcal{F}. \quad (14)$$

*Proof.* In view of Proposition 3 we have  $H_\nu = U_\nu H_\nu = \sum_{i \in \mathcal{R}_\nu} \Delta_i H_\nu$  and, particularly,  $\Delta_i H_\nu = 0$  for  $i \notin \mathcal{R}_\nu$ , since  $H_\nu \in \mathcal{P}_{\mathcal{R}_\nu}$ . Therefore

$$(I - U_\Lambda)H_\nu = \sum_{i \in \mathcal{R}_\nu} \Delta_i H_\nu - \sum_{i \in \Lambda} \Delta_i H_\nu = \sum_{i \in \mathcal{R}_\nu} \Delta_i H_\nu - \sum_{i \in \Lambda \cap \mathcal{R}_\nu} \Delta_i H_\nu = \sum_{i \in \mathcal{R}_\nu \setminus \Lambda} \Delta_i H_\nu,$$

giving

$$c_\nu = \max_{\Lambda \subseteq \mathcal{R}_\nu} \|(I - U_\Lambda)H_\nu\|_{L_\mu^2} \leq \max_{\Lambda \subseteq \mathcal{R}_\nu} \sum_{i \in \mathcal{R}_\nu \setminus \Lambda} \|\Delta_i H_\nu\|_{L_\mu^2} \leq \sum_{i \in \mathcal{R}_\nu} \|\Delta_i H_\nu\|_{L_\mu^2}.$$

Moreover, if (13) holds, then

$$\begin{aligned} c_\nu &\leq \sum_{i \in \mathcal{R}_\nu} \|\Delta_i H_\nu\|_{L_\mu^2} = \sum_{i \in \mathcal{R}_\nu} \prod_{m \in \mathbb{N}} \|\Delta_{i_m} H_{\nu_m}\|_{L_\mu^2} \leq \sum_{i \in \mathcal{R}_\nu} \prod_{m \in \mathbb{N}} (1 + K\nu_m)^\theta \\ &= |\mathcal{R}_\nu| \prod_{m \in \mathbb{N}} (1 + K\nu_m)^\theta \leq \prod_{m \in \mathbb{N}} (1 + K\nu_m)^{\theta+1}. \end{aligned}$$

where we have used (10) and  $K \geq 1$  in the last inequality.  $\square$

**Remark 5.** Bounds such as (13) can often be found in the sparse collocation or sparse quadrature literature, e.g., for quadrature operators applied to Hermite polynomials [9], norms of quadrature operators on bounded domains [35] or Lebesgue constants for Leja points [12]. Numerical estimates for the specific case of Genz-Keister points have been provided in [7].

The following lemma provides a natural starting point for bounding the approximation error of  $U_\Lambda f$  for monotone subsets  $\Lambda$ . The proof follows the same line of arguments as in the proof of [9, Lemma 3.2].

**Lemma 6** (cf. [9, Lemma 3.2]). For a finite and monotone subset  $\Lambda \subset \mathcal{F}$  there holds

$$\|f - U_\Lambda f\|_{L_\mu^2} \leq \sum_{\nu \in \mathcal{F} \setminus \Lambda} c_\nu \|f_\nu\|_{\mathcal{H}}. \quad (15)$$

*Proof.* Due to the monotonicity of  $\Lambda$  we can apply Proposition 3 and obtain

$$\begin{aligned} \|f - U_\Lambda f\|_{L_\mu^2} &= \left\| \sum_{\nu \in \mathcal{F}} f_\nu (I - U_\Lambda) H_\nu(\xi) \right\|_{L_\mu^2} = \left\| \sum_{\nu \in \mathcal{F} \setminus \Lambda} f_\nu (I - U_\Lambda) H_\nu(\xi) \right\|_{L_\mu^2} \\ &\leq \sum_{\nu \in \mathcal{F} \setminus \Lambda} \|f_\nu\|_{\mathcal{H}} \|(I - U_\Lambda) H_\nu\|_{L_\mu^2} \leq \sum_{\nu \in \mathcal{F} \setminus \Lambda} c_\nu \|f_\nu\|_{\mathcal{H}}. \end{aligned}$$

□

Building on Lemma 6 the approximation error  $\|f - U_\Lambda f\|_{L_\mu^2}$  may be further bounded given summability results for the sequence  $(c_\nu \|f_\nu\|_{\mathcal{H}})_{\nu \in \mathcal{F}}$ . The key result here is known as *Stechkin's lemma* which provides a decay rate for the  $\ell^q$ -tail of an  $\ell^p$ -summable sequence for  $q > p$  and is due to Stechkin for  $q = 2$ .

**Lemma 7** (Stechkin's lemma [13, Lemma 3.6]). Let  $0 < p < q < \infty$  and let

$$(a_\nu)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F}) := \left\{ (b_\nu)_{\nu \in \mathcal{F}} : \sum_{\nu \in \mathcal{F}} |b_\nu|^p < \infty \right\}$$

be a sequence of nonnegative numbers. Then for  $\Lambda_N$  denoting the set of multi-indices  $\nu$  corresponding to the  $N$  largest elements  $a_\nu$ , there holds

$$\left( \sum_{\nu \notin \Lambda_N} a_\nu^q \right)^{1/q} \leq \|(a_\nu)_{\nu \in \mathcal{F}}\|_{\ell^p} (N+1)^{-s}, \quad s = \frac{1}{p} - \frac{1}{q}. \quad (16)$$

The index sets  $\Lambda_N$  in Stechkin's lemma associated with the  $N$  largest sequence elements are not necessarily monotone and, therefore Lemma 6 and Lemma 7 can not be combined to bound the error without additional assumptions. An obvious way to ensure monotonicity of the sets  $\Lambda_N$  in Stechkin's lemma is to assume the sequence  $(a_\nu)$  to be *nonincreasing*, i.e.,

$$\nu \leq \tilde{\nu} \quad \Rightarrow \quad a_\nu \geq a_{\tilde{\nu}}.$$

This leads to

**Theorem 8.** Let Assumptions A1 and A2 be satisfied and let there exist a nonincreasing sequence  $(\hat{c}_\nu)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$  with  $p \in (0, 1)$  such that

$$c_\nu \|f_\nu\|_{\mathcal{H}} \leq \hat{c}_\nu \quad \forall \nu \in \mathcal{F}.$$

Then there exists a nested sequence  $(\Lambda_N)_{N \in \mathbb{N}}$  of finite and monotone subsets  $\Lambda_N \subset \mathcal{F}$  with  $|\Lambda_N| = N$  such that (11) holds with rate  $s = 1/p - 1$ .



We will provide a proof below. The convergence analysis in [12, 35] for sparse quadrature and interpolation in case of a bounded  $\Gamma$  follows Theorem 8, although sometimes hidden in the details. There the authors employ explicit bounds on the norms of the Legendre or Taylor coefficients of  $f : \Gamma \rightarrow \mathcal{H}$  to construct a dominating and nonincreasing sequence  $(\hat{c}_\nu)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$ ,  $p \in (0, 1)$ .

In our setting it is however not always possible to derive explicit bounds on the norm of the Hermite coefficients  $\|f\|_{\mathcal{H}}$ . In [4] a technique was developed which relies on somewhat implicit bounds on  $\|f\|_{\mathcal{H}}$  via a weighted  $\ell^2$ -summability property. We state this approach in

**Theorem 9.** Let Assumptions A1 and A2 be satisfied and let there exist a sequence  $(b_\nu)_{\nu \in \mathcal{F}}$  of positive numbers such that

$$\sum_{\nu \in \mathcal{F}} b_\nu \|f_\nu\|_{\mathcal{H}}^2 < \infty \quad (17)$$

and another sequence  $(\hat{c}_\nu)_{\nu \in \mathcal{F}}$  which is nonincreasing, belongs to  $\ell^p(\mathcal{F})$  for some  $p \in (0, 2)$  and dominates

$$\frac{c_\nu}{b_\nu^{1/2}} \leq \hat{c}_\nu \quad \forall \nu \in \mathcal{F}.$$

Then there exists a nested sequence  $(\Lambda_N)_{N \in \mathbb{N}}$  of finite and monotone subsets  $\Lambda_N \subset \mathcal{F}$  with  $|\Lambda_N| = N$  such that (11) holds with rate  $s = 1/p - 1/2$ .

*Proof of Theorem 8 and Theorem 9.* Let  $\Lambda_N$  be the set of multi-indices  $\nu$  corresponding to the  $N$  largest elements of  $(\hat{c}_\nu)_{\nu \in \mathcal{F}}$ . Then each  $\Lambda_N$  is monotone and the sequence  $(\Lambda_N)_{N \in \mathbb{N}}$  can be chosen to be nested.

If the assumption of Theorem 8 hold, we can apply Lemma 6 and Stechkin's lemma with  $q = 1 > p$  to obtain

$$\|f - U_{\Lambda_N} f\|_{L_\mu^2} \leq \sum_{\nu \in \mathcal{F} \setminus \Lambda_N} c_\nu \|f_\nu\|_{\mathcal{H}} \leq \sum_{\nu \in \mathcal{F} \setminus \Lambda_N} \hat{c}_\nu \leq C(N+1)^{-(1/p-1)}$$

where  $C = \|(\hat{c}_\nu)_{\nu \in \mathcal{F}}\|_{\ell^p}$ .

If the assumptions of Theorem 9 hold, Lemma 6 combined with the Cauchy-Schwartz inequality and Stechkin's lemma for  $q = 2 > p$  give

$$\begin{aligned} \|f - U_{\Lambda_N} f\|_{L_\mu^2} &\leq \sum_{\nu \in \mathcal{F} \setminus \Lambda_N} c_\nu \|f_\nu\|_{\mathcal{H}} = \sum_{\nu \in \mathcal{F} \setminus \Lambda_N} \left( \frac{c_\nu}{b_\nu^{1/2}} \right) (b_\nu^{1/2} \|f_\nu\|_{\mathcal{H}}) \\ &\leq \left( \sum_{\nu \in \mathcal{F} \setminus \Lambda_N} b_\nu \|f_\nu\|_{\mathcal{H}}^2 \right)^{1/2} \cdot \left( \sum_{\nu \in \mathcal{F} \setminus \Lambda_N} \frac{c_\nu^2}{b_\nu} \right)^{1/2} \\ &\leq \left( \sum_{\nu \in \mathcal{F}} b_\nu \|f_\nu\|_{\mathcal{H}}^2 \right)^{1/2} \cdot \left( \sum_{\nu \in \mathcal{F} \setminus \Lambda_N} \hat{c}_\nu^2 \right)^{1/2} \\ &\leq C(N+1)^{-(1/p-1/2)}, \end{aligned}$$

where now  $C = \|(b_\nu^{1/2} \|f_\nu\|)_{\nu \in \mathcal{F}}\|_{\ell^2} \cdot \|(\hat{c}_\nu)_{\nu \in \mathcal{F}}\|_{\ell^p}$ , respectively.  $\square$

The analysis in [9] for sparse quadrature follows the reasoning of Theorem 9.

**Remark 10.** We mention that sparse collocation attains a smaller convergence rate than best  $N$ -term approximation in case the assumptions of Theorem 9 hold. Namely, under these assumptions the best- $N$ -term rate is  $s = \frac{1}{p}$ , see [4, Theorem 1.2]. This reduced convergence rate is not caused by the additional factors

$c_\nu$  in the error analysis of sparse collocation. The reason for the slower rate is missing orthogonality: in the proof of Lemma 6 we could not apply Parseval's identity and had to use the triangle inequality to bound the error. This led to bounds in terms of  $\|f_\nu\|_{\mathcal{H}}$  rather than  $\|u_\nu\|_{\mathcal{H}}^2$  as in the case of orthogonal projections like a best- $N$ -term approximation.

At this point we would like to highlight the importance and nontriviality of constructing a dominating sequence which is nonincreasing. Without the latter property we can not conclude that Stechkin's lemma yields monotone multiindex sets  $\Lambda_N$ , which in turn prohibits to use Lemma 6 as the starting point of our error analysis. Of course, we could consider monotone envelopes  $\Lambda_N \subset \tilde{\Lambda}_N$  of  $\Lambda_N$ , but their size might grow quite fast with  $N$  (e.g., polynomially or even faster, see counterexample below). Moreover, it is not at all obvious that for a sequence  $(a_\nu) \in \ell^p(\mathcal{F})$  there exists a dominating and nonincreasing  $(\hat{a}_\nu) \in \ell^p(\mathcal{F})$ . In particular, we provide the following counterexample: let  $\mathcal{F} = \mathbb{N}$  for simplicity and define  $a_n, n \in \mathbb{N}$  by

$$a_n = \begin{cases} \frac{1}{m^2}, & n = \sum_{k=1}^m k, \\ 0, & \text{otherwise,} \end{cases}$$

i.e.,  $a_1 = 1, a_2 = 0, a_3 = \frac{1}{4}, a_4 = 0, a_5 = 0, a_6 = \frac{1}{9}, a_7 = 0, \dots, a_9 = 0, a_{10} = \frac{1}{16}, a_{11} = 0, \dots$ . Then  $(a_n)_{n \in \mathbb{N}} \in \ell^1(\mathbb{N})$ . The smallest positive nonincreasing sequence  $(\hat{a}_n)$  for any  $(a_n) \in \mathbb{R}^{\mathbb{N}}$  is defined by  $\hat{a}_n := \sup_{m \geq n} |a_m|$ , see [13, Section 3.8]. In our case, we get

$$\hat{a}_n = \frac{1}{m^2} \quad \text{for each } n \text{ such that} \quad 1 + \sum_{k=1}^{m-1} k \leq n \leq \sum_{k=1}^m k$$

and, thus,

$$\sum_{n=1}^{\infty} |\hat{a}_n| = \sum_{m=1}^{\infty} m \frac{1}{m^2} = \infty.$$

Although the example is rather pathological it illustrates that for  $(a_\nu) \in \ell^p(\mathcal{F})$  there need not exist an  $\ell^p$ -summable nonincreasing dominating sequence.

### 3.2 Sufficient Conditions for Weighted Summability and Majorization

We will now follow the strategy of Theorem 9 and study under which requirements the assumptions of Theorem 9 hold. Thereto we recall a result from [4] for weighted  $\ell^2$ -summability of Hermite coefficients  $\|f_\nu\|_{\mathcal{H}}$  given the following smoothness conditions on  $f$ :

**Assumption A3.** Let  $f$  satisfy Assumption A1. There exists an integer  $r \in \mathbb{N}_0$  and a sequence of positive numbers  $(\tau_m^{-1})_{m \in \mathbb{N}} \in \ell^p(\mathbb{N})$ ,  $p > 0$ , such that

- (a) for any  $\alpha \in \mathcal{F}$  with  $|\alpha|_\infty \leq r$  the partial derivative  $\partial^\alpha f : \Gamma \rightarrow \mathcal{H}$ , exists and satisfies  $\partial^\alpha f \in L^2(\mathbb{R}^{\mathbb{N}}; \mathcal{H})$  where, for  $\alpha \in \mathcal{F}$  with  $\alpha_m = 0$  for  $m > n$  we denote by  $\partial^\alpha f$  the derivative  $\partial_{\xi_1}^{\alpha_1} \cdots \partial_{\xi_n}^{\alpha_n} f$ ,
- (b) there holds

$$\sum_{|\alpha|_\infty \leq r} \frac{\tau^{2\alpha}}{\alpha!} \int_{\Gamma} \|\partial^\alpha f(\xi)\|_{\mathcal{H}}^2 d\mu < \infty, \quad (18)$$

where  $\tau^\alpha = \prod_{m=1}^{\infty} \tau_m^{\alpha_m}$  and  $\alpha! = \prod_{m=1}^{\infty} \alpha_m!$ .

Observe that the sum in (18) is actually a series, because  $\alpha$  has infinite many components and therefore there are countably many vectors such that  $|\alpha|_\infty \leq r$ . Assumption A3(a) states that we require a *finite* order of partial differentiability of  $f$ , i.e., up to order  $r$  with respect to each variable  $\xi_m$ , and, maybe more

importantly, Assumption A3(b) asks for a *weighted square-summability* of the  $L_\mu^2$ -norms of the corresponding partial derivatives. The latter, in particular, implies bounds of the form

$$\|\partial^\alpha f\|_{L_\mu^2} \leq K\sqrt{\alpha!} \tau^{-\alpha}, \quad |\alpha|_0 \leq r,$$

since otherwise the summability requirement (18) would not hold. Recalling that  $(\tau_m^{-1})_{m \in \mathbb{N}} \in \ell^p(\mathbb{N})$  this bound implies that, e.g., the  $L_\mu^2$ -norm of the derivative  $\partial_{\xi_m}^\alpha f$ ,  $\alpha \leq r$ , decays if  $m \rightarrow \infty$ .

**Theorem 11** (cf. [4, Theorem 3.1]). Let Assumption A3 be satisfied. Then, with the weights

$$b_\nu = b_\nu(\tau, r) = \sum_{|\alpha|_\infty \leq r} \binom{\nu}{\alpha} \tau^{2\alpha} = \prod_{m \geq 1} \left( \sum_{l=0}^r \binom{\nu_m}{l} \tau_m^{2l} \right), \quad \nu \in \mathcal{F}, \quad (19)$$

where

$$\binom{\nu}{\alpha} := \prod_{m \geq 1} \binom{\nu_m}{\alpha_m} \quad \text{and} \quad \binom{\nu_m}{\alpha_m} := 0 \quad \text{if} \quad \alpha_m > \nu_m,$$

there holds

$$\sum_{\nu \in \mathcal{F}} b_\nu \|f_\nu\|_{\mathcal{H}}^2 = \sum_{|\alpha|_\infty \leq r} \frac{\tau^{2\alpha}}{\alpha!} \int_{\Gamma} \|\partial^\alpha f(\xi)\|_{\mathcal{H}}^2 d\mu < \infty. \quad (20)$$

In [4] the statement of Theorem 11 was actually proven without requiring that both series in (20) be finite. Thus, under the smoothness condition given in Assumption A3 we can ensure the first assumption of Theorem 9. It remains to prove the existence of a nonincreasing and  $\ell^p$ -summable sequence which dominates  $\frac{c_\nu}{b_\nu^{1/2}}$ ,  $\nu \in \mathcal{F}$ . Since the  $b_\nu$  are explicitly given in (19), this boils down to the question, how fast the projection errors  $c_\nu$  are allowed to grow. As it turns out, a polynomial growth w.r.t.  $\nu$  as given in (14) in Proposition 4 is sufficient. We therefore state the following lemma.

**Lemma 12.** Let there exists a  $\theta \in \mathbb{R}_+$  and a  $K \geq 1$  such that

$$c_\nu \leq \prod_{m \geq 1} (1 + K\nu_m)^{\theta+1}, \quad \nu \in \mathcal{F}.$$

Then for any increasing sequence  $(\tau_m)_{m \in \mathbb{N}}$  such that  $\sum_{m \geq 1} \tau_m^{-p} < \infty$  for a  $p > 0$  and for any  $r > 2(\theta + 1) + \frac{2}{p}$  there exists a nonincreasing sequence  $(\hat{c}_\nu)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$  such that

$$\frac{c_\nu}{b_\nu^{1/2}} \leq \hat{c}_\nu \quad \forall \nu \in \mathcal{F},$$

where  $b_\nu = b_\nu(\tau, r)$  is as in (19).

This result is an extension of [9, Lemma 3.4] (which itself is strongly based on [4, Lemma 5.1]) in that here we allow for general  $K$  and  $\theta$ , and we also prove the sequence  $(\hat{c}_\nu)_{\nu \in \mathcal{F}}$  is nonincreasing.

*Proof.* We start with constructing the dominating sequence  $(\hat{c}_\nu)_{\nu \in \mathcal{F}}$  and show afterwards that it belongs to  $\ell^p(\mathcal{F})$  and is nonincreasing. In the following we use the notation  $a \wedge b := \min(a, b)$  and  $a \vee b := \max(a, b)$ .

**Step 1: Constructing  $\hat{c}_\nu$ .** We get due to  $\binom{\nu_m}{\nu_m \wedge r} \tau_m^{2(\nu_m \wedge r)} \leq \binom{\nu_m}{r} \tau_m^{2r} \leq \sum_{l=0}^r \binom{\nu_m}{l} \tau_m^{2l}$  that

$$\frac{c_\nu^2}{b_\nu} \leq \prod_{m \geq 1} \frac{(1 + K\nu_m)^{2(\theta+1)}}{\sum_{l=0}^r \binom{\nu_m}{l} \tau_m^{2l}} \leq \prod_{m \geq 1} \frac{(1 + K\nu_m)^{2\theta+2}}{\binom{\nu_m}{\nu_m \wedge r} \tau_m^{2(\nu_m \wedge r)}} =: \prod_{m \geq 1} \tau_m^{-2(\nu_m \wedge r)} h(\nu_m) \quad (21)$$

where we defined the auxiliary function  $h(n) := \frac{(1+Kn)^{2\theta+2}}{\binom{n}{n \wedge r}}$ ,  $n \in \mathbb{N}$ . We will now derive bounds for  $h(n)$  as well as for  $\tau_m^{-2(\nu_m \wedge r)}$  in order to construct a dominating sequence  $\hat{c}_\nu$ .

For  $n \leq r$  we get  $h(n) = (1 + Kn)^{2\theta+2}$ , but for  $n > r$  holds

$$h(n) = \frac{(1 + Kn)^{2\theta+2}}{\binom{n}{r}} = \frac{r! (1 + Kn)^{2\theta+2}}{(n+1) \cdots (n+r)}.$$

Thus, we have  $h \in \mathcal{O}(n^{2\theta+2-r})$ , i.e., there exists a  $1 \leq C_h < \infty$  such that

$$h(n) \leq C_h n^{2\theta+2-r} =: \hat{h}(n) \quad \forall n \in \mathbb{N}.$$

By setting  $\hat{h}(0) = 1 = h(0)$ , we get  $h(n) \leq \hat{h}(n)$  for all  $n \in \mathbb{N}_0$ .

Furthermore, since  $(\tau_m^{-1})_{m \in \mathbb{N}} \in \ell^p(\mathbb{N})$  we have  $\tau_m \rightarrow \infty$  as  $m \rightarrow \infty$ . Thus, there exists an  $M \in \mathbb{N}$  such that  $\tau_m \geq \sqrt{C_h}$  for  $m \geq M$  and  $\tau_m \leq \sqrt{C_h}$  for  $m < M$ . We define

$$\hat{\tau}_m := \sqrt{C_h} \vee \tau_m, \quad m \in \mathbb{N},$$

and notice that  $\hat{\tau}_m \geq 1$  as well as  $(\hat{\tau}_m^{-1})_{m \in \mathbb{N}} \in \ell^p(\mathbb{N})$  by assumption. Moreover, we obtain for  $m \geq M$

$$\tau_m^{2(\nu_m \wedge r)} = \hat{\tau}_m^{2(\nu_m \wedge r)} \geq \hat{\tau}_m^{2(\nu_m \wedge 1)}, \quad \forall \nu_m \in \mathbb{N}_0,$$

since  $\tau_m = \hat{\tau}_m \geq \sqrt{C_h} \geq 1$  in this case. Further, let us define

$$C_\tau := \min_{m \geq 1} \min_{n=0, \dots, r} \frac{\tau_m^{2n}}{C_h^{n \wedge 1}} > 0$$

which then yields for  $1 \leq m < M$

$$\tau_m^{2(\nu_m \wedge r)} \geq C_\tau C_h^{\nu_m \wedge 1} = C_\tau \hat{\tau}_m^{2(\nu_m \wedge 1)}, \quad \forall \nu_m \in \mathbb{N}_0$$

since  $\hat{\tau}_m = \sqrt{C_h}$  for  $m < M$ .

We now define

$$\hat{c}_\nu^2 := C_\tau^{-M} \prod_{m \geq 1} \hat{\tau}_m^{-2(\nu_m \wedge 1)} \hat{h}(\nu_m). \quad (22)$$

and notice that  $\hat{c}_\nu^2$  dominates  $\frac{c_\nu^2}{b_\nu}$  by (21).

**Step 2: Show that  $(\hat{c}_\nu)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$ .** As for the  $p$ -summability, there holds

$$\begin{aligned} \sum_{\nu \in \mathcal{F}} \hat{c}_\nu^p &= C_\tau^{-pM/2} \sum_{\nu \in \mathcal{F}} \prod_{m \geq 1} \hat{\tau}_m^{-p(\nu_m \wedge 1)} \hat{h}^{p/2}(\nu_m) \\ &= C_\tau^{-pM/2} \prod_{m \geq 1} \sum_{n \geq 0} \hat{\tau}_m^{-p(n \wedge 1)} \hat{h}^{p/2}(n). \end{aligned}$$

We get

$$\sum_{n \geq 0} \hat{\tau}_m^{-p(n \wedge 1)} \hat{h}^{p/2}(n) = 1 + C_h^{p/2} \hat{\tau}_m^{-p} \underbrace{\sum_{n \geq 1} n^{-p(r-2\theta-2)/2}}_{=:S}$$

where the sum  $S$  is finite due to the assumption  $\frac{p}{2}(r-2\theta-2) = \frac{p}{2}(r-2\theta-2) > 1$ . The rest follows as in [9, Lemma 3.4] by using  $\log(1+x) \leq x$  for  $x$  positive in order to get

$$\sum_{\nu \in \mathcal{F}} \hat{c}_\nu^p = C_\tau^{-pM/2} \prod_{m \geq 1} (1 + C_h^{p/2} S \hat{\tau}_m^{-p}) \leq C_\tau^{-pM/2} \exp \left( C_h^{p/2} S \sum_{m \geq 1} \hat{\tau}_m^{-p} \right) < \infty$$

since  $(\hat{\tau}_m^{-1})_{m \in \mathbb{N}}$  is in  $\ell^p(\mathbb{N})$  by construction.

**Step 3: Show that  $(\hat{c}_\nu)_{\nu \in \mathcal{F}}$  is nonincreasing.** Let  $\nu \in \mathcal{F}$  be arbitrary. If  $m \in \text{supp } \nu = \{m \in \mathbb{N} : \nu_m > 0\}$ , then we get

$$\hat{c}_{\nu+e_m}^2 = \hat{c}_\nu^2 \cdot \frac{\hat{h}(\nu_m + 1)}{\hat{h}(\nu_m)} \leq \hat{c}_\nu^2,$$

since  $\hat{h}(n)$  is nonincreasing for  $n \geq 1$ . Let now  $m \notin \text{supp } \nu$ . Then

$$\hat{c}_{\nu+e_m}^2 = \hat{c}_\nu^2 \cdot \hat{\tau}_m^{-2} \cdot \hat{h}(1) = \hat{c}_\nu^2 \cdot C_h \hat{\tau}_m^{-2} \leq \hat{c}_\nu^2 \cdot C_h (\sqrt{C_h})^{-2} \leq \hat{c}_\nu^2.$$

In summary, we obtain

$$\hat{c}_{\nu+e_m} \leq \hat{c}_\nu \quad \forall m \in \mathbb{N},$$

hence,  $(\hat{c}_\nu)_{\nu \in \mathcal{F}}$  is nonincreasing.  $\square$

Thus, Lemma 12 provides sufficient conditions for the second assumption of Theorem 9 to hold. Together with Theorem 11 which ensured the first assumption of Theorem 9 and Proposition 4 to get the polynomial growth of the  $c_\nu$  we can now state our main convergence result for sparse collocation.

**Theorem 13** (Convergence of sparse collocation). Assume there holds for a  $\theta \geq 0$  and a  $K \geq 1$

$$\|\Delta_i H_\nu\|_{L_\mu^2} \leq (1 + K\nu)^\theta, \quad i \in \mathbb{N}_0. \quad (23)$$

Then, for any function  $f$  which satisfies Assumption A3 with  $r > 2(\theta + 1) + \frac{2}{p}$  and Assumption A2, there exists a nested sequence of monotone finite subsets  $\Lambda_N \subset \mathcal{F}$  with  $|\Lambda_N| = N$  such that for the sparse collocation error holds

$$\|f - U_{\Lambda_N} f\|_{L_\mu^2} \leq C(1 + N)^{-\left(\frac{1}{p} - \frac{1}{2}\right)}.$$

### 3.3 Convergence of Sparse Collocation Using Gauss-Hermite Nodes

In the following, we will verify the assumption (23) in Corollary 13 for the interpolation operators  $U_i$  based on Gauss-Hermite nodes and provide a convergence result in terms of number of grid points  $|\Xi_{\Lambda_N}|$  rather than in terms of  $N = |\Lambda_N|$ . The latter is not obvious, since Gauss-Hermite nodes are non-nested. Otherwise, for nested univariate node sets, i.e.,  $\Xi_{i+i} = \Xi_i \cup \{\xi_{i+1}^{(i+1)}\}$  we have  $|\Xi_{\Lambda_N}| = |\Lambda_N|$ .

**Lemma 14.** For  $U_i$  being the interpolation operator based on the zeros of the  $(i + 1)$ th Hermite polynomial we have for each  $\nu \in \mathbb{N}$  that

$$\|U_i H_\nu\|_{L_\mu^2}^2 \leq c^2 e^{\sqrt{2\nu - 1}} \quad \forall i \in \mathbb{N}_0$$

where  $c = 1.086435$  is the constant appearing in *Cramér's inequality* for Hermite functions. In particular, there holds

$$\|\Delta_i H_\nu\|_{L_\mu^2} \leq (1 + K\nu)$$

with  $K = 2c\sqrt{e} > 1$ .

*Proof.* We start by recalling the  $L_\rho^2$ -orthogonality ( $\rho$  denoting the probability density function of the standard Gaussian measure  $N(0, 1)$ ) of Lagrange basis polynomials  $L_k^{(i)}$  constructed from the zeros  $\{\xi_k^{(i)}\}_{k=0}^i$  of the Hermite polynomial of degree  $i + 1$  ([37, Theorem 14.2.1]). This orthogonality yields

$$\begin{aligned} \|U_i H_\nu\|_{L_\mu^2}^2 &= \int_{\mathbb{R}} \left( \sum_{k=0}^i H_\nu(\xi_k^{(i)}) L_k^{(i)}(\xi) \right)^2 \rho(\xi) d\xi \\ &= \sum_{k=0}^i H_\nu^2(\xi_k^{(i)}) \int_{\mathbb{R}} \left( L_k^{(i)}(\xi) \right)^2 \rho(\xi) d\xi = \sum_{k=0}^i H_\nu^2(\xi_k^{(i)}) w_k^{(i)} \end{aligned}$$

where  $\{w_k^{(i)}\}_{k=0}^i$  denotes the weights of the Gauss quadrature formulae based on the zeros of the  $(i + 1)$ th Hermite polynomial, see also [37, Theorem 14.2.1].

Next, we recall Cramér's inequality for the Hermite polynomials  $\tilde{H}_\nu$  taken w.r.t. the weight  $\tilde{\rho}(\xi) = \exp(-\xi^2)$ , i.e.,

$$|\tilde{H}_n(\xi)| \leq c\pi^{-1/4} \exp(\xi^2/2),$$

see, e.g., [1, Chapter 22, p.787]. Since there holds  $\tilde{H}_n(\xi) = \pi^{-1/4} H_n(\xi\sqrt{2})$  [1, Chapter 22, p.778], we get

$$|H_n(\xi)| \leq c \exp(\xi^2/4)$$

and, thus,

$$\|U_i H_\nu\|_{L_\mu^2}^2 \leq c^2 \sum_{k=0}^i \exp(\xi_{ki}^2/2) w_{ki},$$

where we switched notation to  $\xi_{ki} := \xi_k^{(i)}$  and  $w_{ki} := w_k^{(i)}$  for convenience. Furthermore, we use a consequence of [30, Lemma 4]. The latter states for  $\tilde{\xi}_{kn}$  being the zeros of  $\tilde{H}_n$  and  $\tilde{w}_{kn}$  the Christoffel numbers of corresponding Gauss-Hermite quadrature (i.e. Gauss-Hermite weights for  $\tilde{\rho}$ ) that

$$\sum_{k=1}^n \tilde{w}_{kn} \exp(\tilde{\xi}_{kn}^2) \leq e \sqrt{\pi(2n + 1)}.$$

It can be easily verified that

$$\xi_{kn} = \sqrt{2} \tilde{\xi}_{kn} \quad \text{and} \quad w_{kn} = \pi^{-1/2} \tilde{w}_{kn}.$$

Hence, we get

$$\sum_{k=0}^i \exp(\xi_{ki}^2/2) w_{ki} \leq e \sqrt{2(i + 1) + 1}$$

and by noticing that for  $i \geq \nu$  we have  $U_i H_\nu = H_\nu$  and, thus,  $\|U_i H_\nu\|_{L_\mu^2}^2 = 1$ , and for  $i = \nu - 1$  we get  $U_i H_\nu \equiv 0$  the first assertion is shown.

For the second statement we notice

$$\|U_i H_\nu\|_{L_\mu^2}^2 \leq c^2 e \nu, \quad \forall i \in \mathbb{N}_0 \forall \nu \geq 1$$

since  $\nu \geq \sqrt{2\nu - 1}$  for  $\nu \geq 1$ . And, because of  $\Delta_i H_0 \equiv 0$  for  $i \geq 1$  and  $\Delta_0 H_0 \equiv H_0$ , we get

$$\|\Delta_i H_\nu\|_{L_\mu^2} \leq 1 + K\nu, \quad \forall i, \nu \in \mathbb{N}_0.$$

□

Thus, interpolation on Gauss-Hermite points satisfies the assumptions of Corollary 13 with  $\theta = 1$  and we obtain

**Theorem 15** (Convergence of sparse collocation, Gauss-Hermite nodes). For any function  $f$  which satisfies Assumption A3 with  $r > 4 + \frac{2}{p}$  and Assumption A2, there exists a nested sequence of mononote finite subsets  $\Lambda_N \subset \mathcal{F}$  with  $|\Lambda_N| = N$  such that for the error of the sparse collocation operator  $U_{\Lambda_N}$  based on Gauss-Hermite nodes holds

$$\|f - U_{\Lambda_N} f\|_{L_\mu^2} \leq C(1 + N)^{-\left(\frac{1}{p} - \frac{1}{2}\right)}.$$

**Remark 16.** In numerical experiments we actually observed for  $\nu = 0, \dots, 39$  that

$$\|U_i H_\nu\|_{L_\mu^2} \leq 1, \quad \forall i \in \mathbb{N}_0,$$

see Figure 1. This would imply

$$\|\Delta_i H_\nu\|_{L_\mu^2} \leq \begin{cases} 1 & \text{if } \nu = 0, \\ 2 & \text{otherwise,} \end{cases} \quad \forall i, \nu \in \mathbb{N}_0.$$

Again, we even observed a smaller bound numerically, see the right plot in Figure 1. However, we have not been able to prove  $\|U_i H_\nu\|_{L_\mu^2} \leq 1$  and the improvement in the statement of Corollary 15 would have been minor, i.e., the assertion would also hold with the same rate for functions  $f : \Gamma \rightarrow \mathcal{H}$  satisfying Assumption A3 with  $r > 2 + \frac{2}{p}$ . Note that similar numerical evidence was presented in [9] for quadrature operators applied to Hermite polynomials. See also [7] for analogous numerical bounds in the case of Genz-Keister points.

**Convergence rate w.r.t. number of nodes** We will now derive bounds for the number of nodes in the sparse grid associated to  $U_\Lambda$ ,

$$\Xi_\Lambda := \bigcup_{i \in \Lambda} \Xi^{(i)}, \quad (24)$$

in terms of the cardinality  $|\Lambda|$  and combine this bound with the result in Corollary 15. But first, let us show that (24) does indeed provide all points where we need to know  $f$  for constructing  $U_\Lambda f$ : we note that for  $i \in \mathcal{F}$

$$\Delta_i f = \left[ \bigotimes_{m \geq 1} (U_{i_m} - U_{i_m - 1}) \right] f = \sum_{k: i_m - k_m \leq 1} (-1)^{|i - k|_1} \left[ \bigotimes_{m \geq 1} U_{k_m} \right] f,$$

hence, for computing  $\Delta_i f$  we need to evaluate  $f$  at

$$\Xi^{(i), \Delta} := \bigcup_{k: i_m - k_m \leq 1} \Xi^{(k)}.$$

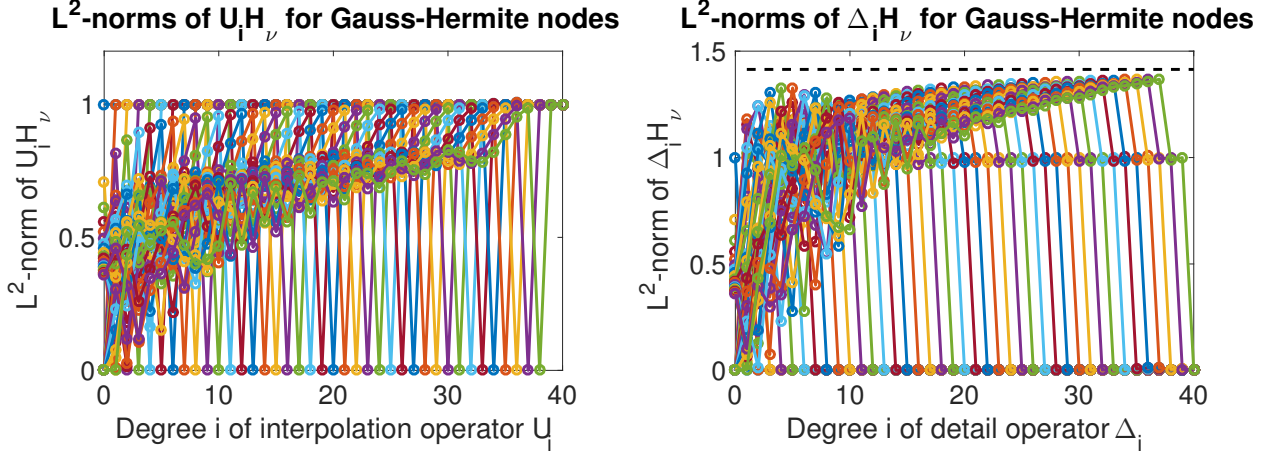


Figure 1: Computed values of  $\|U_i H_\nu\|_{L_\mu^2}$  (left) and  $\|\Delta_i H_\nu\|_{L_\mu^2}$  (right) for Gauss-Hermite nodes. The dashed, black line in the right plot indicates the value  $\sqrt{2}$ .

If  $\Lambda$  is monotone set, then the resulting grid for  $U_\Lambda = \sum_{i \in \Lambda} \Delta_i$  is

$$\bigcup_{i \in \Lambda} \Xi^{(i), \Delta} = \bigcup_{i \in \Lambda} \bigcup_{i-k \leq 1} \Xi^{(k)} = \bigcup_{i \in \Lambda} \Xi^{(i)} = \Xi_\Lambda.$$

Now, consider the following simple monotone index of cardinality  $N$ :  $\Lambda_N = \{0e_j, \dots, (N-1)e_j\}$ . Then due to  $|\Xi_{\{ke_j\}}| = (k+1)$  we get for this  $\Lambda_N$  that

$$|\Xi_{\Lambda_N}| \leq \sum_{k=0}^{N-1} (k+1) = \frac{N(N+1)}{2} \in \mathcal{O}(N^2).$$

The quadratic complexity is indeed sharp, since 0 is the only reappearing Gauss-Hermite node. We will show in the subsequent two propositions that this complexity holds also for arbitrary monotone multiindex sets. We start with rectangular envelopes  $\Lambda = \mathcal{R}_\nu$  and provide also a rather technical ordering result which we will require later on. Recall that  $|\Xi^{(i)}| = \prod_{m \geq 1} (1 + i_m)$ .

**Proposition 17.** Let  $\nu \in \mathcal{F}$ . Then there exists an ordering  $n$  of  $\mathcal{R}_\nu$ , i.e., a bijective mapping  $n : \mathcal{R}_\nu \rightarrow \{1, \dots, |\mathcal{R}_\nu|\}$  such that

$$|\Xi^{(i)}| \leq n(i) \quad \forall i \in \mathcal{R}_\nu,$$

which implies, in particular,

$$|\Xi_{\mathcal{R}_\nu}| \leq \frac{|\mathcal{R}_\nu| (|\mathcal{R}_\nu| + 1)}{2}.$$

*Proof.* The second assertion follows from the first easily by

$$|\Xi_{\mathcal{R}_\nu}| \leq \sum_{i \leq \nu} |\Xi^{(i)}| \leq \sum_{n=1}^{|\mathcal{R}_\nu|} n = \frac{|\mathcal{R}_\nu| (|\mathcal{R}_\nu| + 1)}{2}.$$

We prove the first assertion by induction. Since  $\nu \in \mathcal{F}$ , there exist only finitely many  $m \in \mathbb{N}$  such that  $\nu_m > 0$ . Without loss of generality we assume that  $\nu_m = 0$  for  $m > M$  where  $M \in \mathbb{N}$ . We now perform an induction over the number  $M$  of non-zero entires in  $\nu$ .



- **base case**  $M = 1$ : The only possible  $\nu \in \mathcal{F}$  are  $\nu = k e_1, k \in \mathbb{N}_0$ , and we have  $R_\nu = \{0 e_1, 1 e_1, \dots, \nu_1 e_1\}$ . The ordering is then simply  $n(i) = i_1 + 1$ . Then

$$|\Xi^{((i_1, 0, \dots))}| = 1 + i_1 = n(i).$$

- **Induction step**: the assertion holds for  $M \geq 1$ . Let  $\nu \in \mathcal{F}$  be such that  $\nu_m = 0$  for  $m \geq M + 2$ . Moreover, let  $n_M$  denote the ordering for  $\mathcal{R}_{\nu - \nu_{M+1} e_{M+1}} = \{i \in \mathcal{R}_\nu : i_{M+1} = 0\}$ , i.e., it holds

$$|\Xi^{(i)}| = \prod_{m=1}^M (1 + i_m) \leq n_M(i) \quad \forall i \in \mathcal{R}_{\nu - \nu_{M+1} e_{M+1}}.$$

For notational convenience, we set  $i_M := (i_1, \dots, i_M, 0, \dots)$  for each  $i \in \mathcal{R}_\nu$  and observe that  $i_M \in \mathcal{R}_{\nu - \nu_{M+1} e_{M+1}}$ . We define the ordering

$$n(i) := i_{M+1} \left( \prod_{m=1}^M (1 + \nu_m) \right) + n_M(i_M), \quad i \in \mathcal{R}_\nu.$$

It is easy to check that  $n : \mathcal{R}_\nu \rightarrow \{1, \dots, |\mathcal{R}_\nu|\}$  is again bijective. Furthermore, we get for each  $i \in \mathcal{R}_\nu$

$$\begin{aligned} |\Xi^{(i)}| &= \prod_{m=1}^{M+1} (1 + i_m) = (i_{M+1} + 1) \prod_{m=1}^M (1 + i_m) \\ &= i_{M+1} \left( \prod_{m=1}^M (1 + i_m) \right) + \prod_{m=1}^M (1 + i_m) \\ &= i_{M+1} \left( \prod_{m=1}^M (1 + i_m) \right) + |\Xi_{i_M}| \\ &\leq i_{M+1} \left( \prod_{m=1}^M (1 + \nu_m) \right) + n_M(i_M) = n(i) \end{aligned}$$

where the last line follows by  $i_m \leq \nu_m$  for all  $m \geq 1$  and the fact that  $i_M \in \mathcal{R}_{\nu - \nu_{M+1} e_{M+1}}$  for  $i \in \mathcal{R}_\nu$ . □

We extend the estimate for  $\Xi_{\mathcal{R}_\nu}$  in the above proposition now to arbitrary finite and monotone index sets  $\Lambda$ :

**Proposition 18.** Let  $\Lambda \subset \mathcal{F}$  be a finite and monotone index set. Then for  $\Xi_\Lambda$  as in (24) there holds

$$|\Xi_\Lambda| \leq \frac{|\Lambda| (|\Lambda| + 1)}{2}. \quad (25)$$

*Proof.* Since  $\Lambda$  is supposed to be monotone, it is a union of rectangular envelopes, i.e., there exist  $n$  indices  $\nu_1, \dots, \nu_n \in \mathcal{F}$  such that

$$\Lambda = \bigcup_{k=1}^n \mathcal{R}_{\nu_k} \quad \text{and} \quad \Xi_\Lambda = \bigcup_{k=1}^n \Xi_{\mathcal{R}_{\nu_k}}.$$

We prove the assertion by induction over  $n$ :

- **base case**  $n = 1$ : The assertion follows by the second statement of Proposition 17.

- **Induction step:** the assertion holds for  $n \geq 1$ . With a slight abuse of notation we set  $\Lambda_n := \bigcup_{k=1}^n \mathcal{R}_{\nu_k}$  and obtain

$$\sum_{i \in \Lambda_n \cup \mathcal{R}_{\nu_{n+1}}} |\Xi^{(i)}| = \sum_{i \in \Lambda_n} |\Xi^{(i)}| + \sum_{i \in \mathcal{R}_{\nu_{n+1}} \setminus \Lambda_n} |\Xi^{(i)}|.$$

Let  $m := |\mathcal{R}_{\nu_{n+1}} \setminus \Lambda_n|$ . The first statement of Proposition 17 now implies

$$\sum_{i \in \mathcal{R}_{\nu_{n+1}} \setminus \Lambda_n} |\Xi^{(i)}| \leq \sum_{k=1+|\mathcal{R}_{\nu_{n+1}}|-|\mathcal{R}_{\nu_{n+1}} \setminus \Lambda_n|}^{|\mathcal{R}_{\nu_{n+1}}|} k \leq \sum_{k=1+|\Lambda_n|}^{|\Lambda_n|+|\mathcal{R}_{\nu_{n+1}} \setminus \Lambda_n|} k$$

where the last inequality is due to  $|\mathcal{R}_{\nu_{n+1}}| \leq |\mathcal{R}_{\nu_{n+1}} \setminus \Lambda_n| + |\Lambda_n|$ . Thus, we get by the induction hypothesis

$$\sum_{i \in \Lambda_n \cup \mathcal{R}_{\nu_{n+1}}} |\Xi^{(i)}| \leq \sum_{k=1}^{|\Lambda_n|} k + \sum_{k=1+|\Lambda_n|}^{|\Lambda_n|+|\mathcal{R}_{\nu_{n+1}} \setminus \Lambda_n|} k = \sum_{k=1}^{|\Lambda_n \cup \mathcal{R}_{\nu_{n+1}}|} k.$$

□

Thus, employing non-nested points such as Gauss-Hermite points, yields indeed a quadratic growth of the number of grid points

$$|\Xi_\Lambda| \in \mathcal{O}(|\Lambda|^2)$$

whereas in the nested case one has

$$|\Xi_\Lambda| = |\Lambda|.$$

We provide some numerical validation of bound (25). More precisely, we consider the following two families of multi-index sets  $\Lambda$  (cf. [6]):

**Total Degree (TD)**  $\Lambda = \Lambda(w, M) = \{\nu \in \mathcal{F} : \sum_{m=1}^M \nu_m \leq w, \nu_m = 0 \text{ for } m > M\}$

**Hyperbolic Cross (HC)**  $\Lambda = \Lambda(w, M) = \{\nu \in \mathcal{F} : \prod_{m=1}^M (\nu_m + 1) \leq w, \nu_m = 0 \text{ for } m > M\},$

In Figure 2 we fix the number of (active) dimensions  $M$  and display the cardinality of  $\Xi_{\Lambda(w, M)}$  for both choices of  $\Lambda(w, M)$  and increasing values of  $w \in \mathbb{N}$ . The plot shows that estimate (25) is valid but slightly pessimistic for the two specific examples considered here.

We finally arrive at the resulting error-cost theorem:

**Theorem 19** (Convergence rate of Gauss-Hermite sparse grid collocation in terms of nodes). For any function  $f$  which satisfies Assumption A3 with  $r > 4 + \frac{2}{p}$  and Assumption A2, there exists a nested sequence of mononote finite subsets  $\Lambda_N \subset \mathcal{F}$  with  $|\Lambda_N| = N^{\frac{1}{p}}$  such that for the error of the sparse collocation operator  $U_{\Lambda_N}$  based on Gauss-Hermite nodes holds

$$\|f - U_{\Lambda_N} f\|_{L^2 \mu} \leq C |\Xi_{\Lambda_N}|^{-\left(\frac{1}{2p} - \frac{1}{4}\right)}$$

where  $C$  depends on  $f$ .

Hence, assume we require an approximation error  $\|f - U_{\Lambda_N} f\|_{L^2 \mu} \leq \varepsilon$ , then we can achieve this accuracy with

$$\text{cost}(\varepsilon) \in \mathcal{O}\left(\varepsilon^{\frac{1}{2p} - \frac{1}{4}}\right) \quad (26)$$

number of function evaluations of  $f$ . In this cost complexity (26) we neglected of course the computational work which is necessary to find the resulting multiindex sets  $\Lambda_N$ . This is a very important issue. Typically, they are constructed employing adaptive algorithms, see [11, 35, 32]. Our result makes no statement about the actual computational work of those.

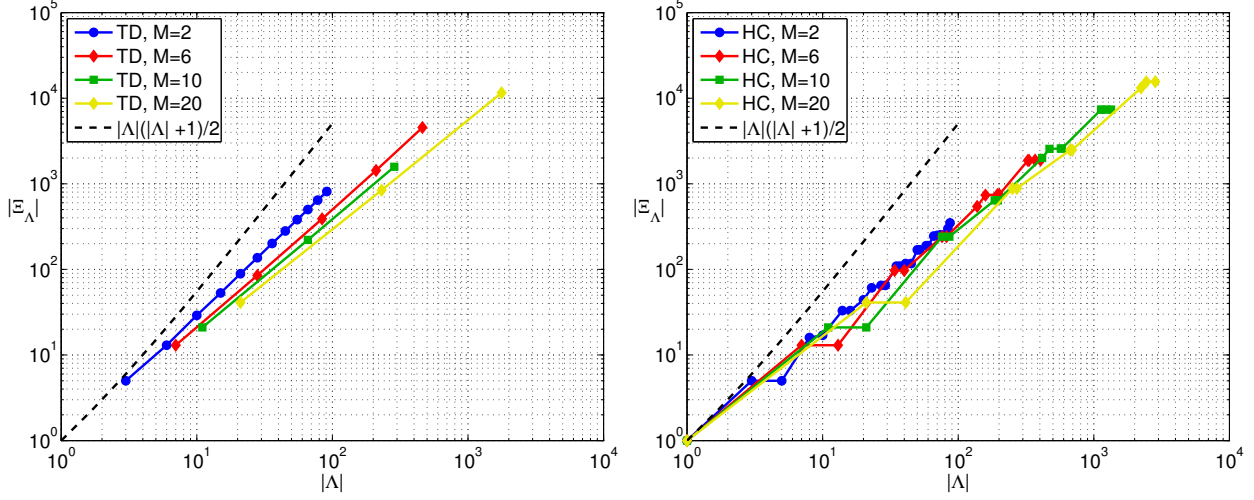


Figure 2: Numerical verification of estimate (25) for “Total Degree” sparse grids (left) and “Hyperbolic Cross” sparse grids (right).

**Remark 20** (On sparse collocation employing weighted Leja points). As mentioned in the introduction weighted Leja points [29] seem to be a promising node family for sparse grid interpolation and approximation. So far, we are however unable to prove bounds like (23) for them. Moreover, we believe that a maybe more suitable approach for analyzing convergence in case of weighted Leja nodes is to measure the approximation error in the weighted  $L^\infty_\mu$ -norm instead of the  $L^2_\mu$ -norm and try to estimate the corresponding Lebesgue constant for weighted Leja nodes. See [25] for first results on the latter – which does not yet imply an analogous estimate to (23) – and [12, 11] for the convergence analysis of sparse grid collocation using Leja points on  $[-1, 1]$  via estimates for the Lebesgue constant of those.

## 4 Application to Elliptic PDEs

We recall our motivation from the introduction: approximating the weak solution  $u$  of an elliptic boundary value problem with lognormal diffusion coefficient as in (1) where  $f \in L^2(D)$  and  $a(\xi) \in L^\infty(D)$  is given as (2). We will discuss now under which conditions the map  $\xi \mapsto u(\xi) \in H_0^1(D)$  satisfies Assumptions A1, A2 and A3 and can therefore be approximated by sparse grid collocation methods based on Gauss-Hermite nodes as outlined in the previous section. We will mainly cite results from [4] but try to emphasize those details which are sometimes omitted in the literature.

**Verifying Assumption A1.** First of all, we have to investigate the domain  $\Gamma$  of the mapping  $\xi \mapsto u(\xi)$ . There holds  $\Gamma \neq \mathbb{R}^N$  since for arbitrary  $\xi \in \mathbb{R}^N$  the expansion (2) need not to converge. A natural domain  $\Gamma$  for the mapping  $\xi \mapsto u(\xi)$  is

$$\Gamma := \left\{ \xi \in \mathbb{R}^N : \left\| \sum_{m=1}^{\infty} \phi_m \xi_m \right\|_{L^\infty(D)} < \infty \right\} \quad (27)$$

which is clearly an open set. Further, a natural condition on the decay of the  $\phi_m$  is

$$\sum_{m=1}^{\infty} \|\phi_m\|_{L^\infty(D)}^2 < \infty \quad (28)$$

since (28) implies that the series (2) converges in  $L_\mu^2(\mathbb{R}^N; L^\infty(D))$ :

$$\begin{aligned} \mathbf{E} \left[ \left\| \sum_{m=1}^{\infty} \phi_m \xi_m \right\|_{L^\infty(D)}^2 \right] &\leq \mathbf{E} \left[ \left( \sum_{m=1}^{\infty} \|\phi_m(x)\|_{L^\infty(D)} |\xi_m| \right)^2 \right] = \sum_{m=1}^{\infty} \|\phi_m(x)\|_{L^\infty(D)}^2 \mathbf{E} [|\xi_m|^2] \\ &= \sum_{m=1}^{\infty} \|\phi_m(x)\|_{L^\infty(D)}^2 \end{aligned}$$

due to  $\mathbf{E} [\xi_m \xi_n] = \delta_{mn}$ . Moreover, by a classical result [26, Lemma 4.16] from probability theory this implies that the series converges also  $\mu$ -a.e. in  $L^\infty(D)$ . Thus, if (28) holds, then we get  $\mu(\Gamma) = 1$ . It remains to state conditions under which we can ensure that  $\xi \mapsto u(\xi)$  belongs to  $L_\mu^2(\Gamma; H_0^1(D))$ . Measurability follows from the continuous dependence of the weak solution  $u \in H_0^1(D)$  on  $\exp(a) \in L^\infty(D)$ , see [21]. Moreover, if we can ensure that for  $\underline{a}(\xi) := \text{essinf}_{x \in D} \exp(a(x, \xi))$  we have  $\underline{a}^{-1} \in L_\mu^2(\Gamma; \mathbb{R})$  (e.g. via the Fernique lemma, as shown in [8]) then the  $\xi$ -pointwise application of the Lax-Milgram lemma [21] yields for the random solution  $u \in L_\mu^2(\Gamma; H_0^1(D))$ . The latter can be guaranteed by an even weaker assumption than (28)

**Assumption A4** ([4, Assumption A]). There exists a strictly positive sequence  $(\rho_m)_{m \in \mathbb{N}}$  such that

$$\sup_{x \in D} \sum_{m=1}^{\infty} \rho_m |\phi_m(x)| < \infty, \quad \sum_{m=1}^{\infty} \exp(-\rho_m^2) < \infty.$$

Then by [4, Corollary 2.1] it follows that if Assumption A4 holds we have  $u \in L_\mu^2(\Gamma; H_0^1(D))$ , and thus,  $u : \Gamma \rightarrow H_0^1(D)$  satisfies Assumption A1.

**Verifying Assumption A2.** It is obvious that for Gauss-Hermite nodes there holds  $\Xi^{(i)} \subset \Gamma$ ,  $i \in \mathcal{F}$ , with  $\Gamma$  as in (27), because due to  $i \in \mathcal{F}$  there exists an  $M \in \mathbb{N}$  such that for  $\xi \in \Xi^{(i)}$  we have  $\xi_m = \xi_0^{(0)}$  for any  $m \geq M$  and  $\xi_0^{(0)} = 0$ . [Note: it is enough to require that  $\xi \in \ell^\infty(\mathbb{N})$ ]

**Verifying Assumption A3.** Again, we refer to results from [4], namely, [4, Theorem 4.2] where the authors show that under

**Assumption A5.** There exists a strictly positive sequence  $(\rho_m^{-1})_{m \in \mathbb{N}} \in \ell^p(\mathbb{N})$  such that

$$\sup_{x \in D} \sum_{m=1}^{\infty} \rho_m |\phi_m(x)| < \infty.$$

the weak solution  $u$  of (1) satisfies Assumption A3 for any  $r \in \mathbb{N}_0$ . Please note that Assumption A5 implies Assumption A4, see [4, Remark 2.2]. Hence, we obtain

**Corollary 21.** Let  $a$  be given as in (2) and satisfy Assumption A5. Then there exists a nested sequence of mononote finite subsets  $\Lambda_N \subset \mathcal{F}$  with  $|\Lambda_N| = N$  such that for the sparse collocation operator  $U_{\Lambda_N}$  based on Gauss-Hermite nodes applied to the weak solution  $u$  of (1) holds

$$\|u - U_{\Lambda_N} u\|_{L_\mu^2(\mathbb{R}^N; H_0^1(D))} \leq C |\Xi_{\Lambda_N}|^{-\left(\frac{1}{2p} - \frac{1}{4}\right)}.$$

## 5 Conclusions

In this note we have presented the general convergence analysis of adaptive sparse grid collocation based on Lagrange interpolation for functions of countably many variables in unbounded domains with Gaussian probability density function as it can be found in many recent works. In particular, we have stated sufficient conditions for the underlying univariate interpolation nodes such that for functions of a certain smoothness there holds an algebraic rate of convergence for the adaptive sparse grid approximation. Moreover, we verified these assumptions for the classical Gauss-Hermite nodes and were able to state also a convergence results in terms of the resulting number of grid nodes in this case by deriving sharp bounds on the latter. We finally discussed in detail that these methods can be applied to weak solutions of lognormal diffusion problems mainly building up on previous work.

## References

- [1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of Mathematical Functions*, vol. 55 of Applied Mathematics Series, National Bureau of Standards, Washington, 10th ed., 1972.
- [2] R. J. ADLER, *The Geometry of Random Fields*, John Wiley & Sons, New York, 1981.
- [3] I. BABUŠKA, F. NOBILE, AND R. TEMPONE, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Review, 52 (2010), pp. 317–355.
- [4] M. BACHMAYR, A. COHEN, R. DEVORE, AND G. MIGLIORATI, *Sparse polynomial approximation of parametric elliptic PDEs. part II: lognormal coefficients*, ESAIM Math. Model. Numer. Anal., (2016).
- [5] M. BACHMAYR, A. COHEN, AND G. MIGLIORATI, *Sparse polynomial approximation of parametric elliptic PDEs. part I: affine coefficients*, ESAIM Math. Model. Numer. Anal., (2016).
- [6] J. BÄCK, F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison*, in Spectral and High Order Methods for Partial Differential Equations, vol. 76 of Lecture Notes in Computational Science and Engineering, Springer, 2011, pp. 43–62.
- [7] J. BECK, F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *A Quasi-optimal Sparse Grids Procedure for Groundwater Flows*, in Spectral and High Order Methods for Partial Differential Equations - ICOSA-HOM 2012, vol. 95 of Lecture Notes in Computational Science and Engineering, Springer, 2014, pp. 1–16.
- [8] J. CHARRIER, *Strong and weak error estimates for elliptic partial differential equations with random coefficients*, SIAM Journal on Numerical Analysis, 50 (2012), pp. 216–246.
- [9] P. CHEN, *Convergence analysis of an adaptive sparse quadrature for high-dimensional integration with Gaussian random variables*. arXiv:1604.08466, 2016.
- [10] A. CHKIFA, *On the Lebesgue constant of Leja sequences for the complex unit disk and of their real projection*, Journal of Approximation Theory, 166 (2013), pp. 176 – 200.
- [11] A. CHKIFA, A. COHEN, AND C. SCHWAB, *High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs*, Foundations of Computational Mathematics, 14 (2014), pp. 601–633.

- [12] ———, *Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs*, J. Math. Pures Appl., 103 (2015), pp. 400–428.
- [13] A. COHEN AND R. DEVORE, *Approximation of high-dimensional parametric PDEs*, Acta Numerica, 24 (2015), pp. 1–159.
- [14] A. COHEN, R. DEVORE, AND C. SCHWAB, *Convergence rates of best  $N$ -term Galerkin approximations for a class of elliptic sPDEs*, Foundations of Computational Mathematics, 10 (2010), pp. 615–646.
- [15] A. COHEN, R. DEVORE, AND C. SCHWAB, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs*, Analysis and Applications, 9 (2011), pp. 11–47.
- [16] S. DE MARCHI, *On Leja sequences: Some results and applications*, Applied Mathematics and Computation, 152 (2004), pp. 621–647.
- [17] J. GALVIS AND M. SARKIS, *Approximating infinity-dimensional stochastic Darcy’s equations without uniform ellipticity*, SIAM J. Numer. Anal., 47 (2009), pp. 3624–3651.
- [18] A. GENZ AND B. D. KEISTER, *Fully symmetric interpolatory rules for multiple integrals over infinite regions with Gaussian weight*, Journal of Computational and Applied Mathematics, 71 (1996), pp. 299–309.
- [19] T. GERSTNER AND M. GRIEBEL, *Dimension-adaptive tensor-product quadrature*, Computing, 71 (2003), pp. 65–87.
- [20] R. GHANEM AND P. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.
- [21] D. GILBARG AND N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin Heidelberg, 2001.
- [22] C. J. GITTELSON, *Stochastic Galerkin discretization of the log-normal isotropic diffusion problem*, Math. Models Methods Appl. Sci., 20 (2010), pp. 237–263.
- [23] A.-L. HAJI-ALI, F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *Multi-index Stochastic Collocation convergence rates for random PDEs with parametric regularity*, Foundations of Computational Mathematics, (2016).
- [24] V. H. HOANG AND C. SCHWAB,  *$N$ -term Wiener chaos approximation rates for elliptic PDEs with lognormal Gaussian random inputs*, Mathematical Models and Methods in Applied Sciences, 24 (2014), pp. 797–826.
- [25] P. JANTSCH, C. G. WEBSTER, AND G. ZHANG, *On the Lebesgue constant of weighted Leja points for Lagrange interpolation on unbounded domains*. arXiv:1606.07093, 2016.
- [26] O. KALLENBERG, *Foundations of Modern Probability*, Springer, New York, 2002.
- [27] O. P. LE MAITRE AND O. M. KNIO, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Scientific Computation, Springer, New York, 2010.
- [28] A. MUGLER AND H.-J. STARKLOFF, *On the convergence of the stochastic Galerkin method for random elliptic partial differential equations*, ESAIM: Mathematical Modelling and Numerical Analysis, 47 (2013), pp. 1237–1263.

- [29] A. NARAYAN AND J. D. JAKEMAN, *Adaptive Leja sparse grid constructions for stochastic collocation and high-dimensional approximation*, SIAM Journal on Scientific Computing, 36 (2014), pp. A2952–A2983.
- [30] P. G. NEVAI, *Mean convergence of Lagrange interpolation, II*, Journal of Approximation Theory, 30 (1980), pp. 263–276.
- [31] F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *Convergence of quasi-optimal sparse-grid approximation of Hilbert-space-valued functions: application to random elliptic PDEs*, Numerische Mathematik, 134 (2016), pp. 343–388.
- [32] F. NOBILE, L. TAMELLINI, F. TESEI, AND R. TEMPONE, *An adaptive sparse grid algorithm for elliptic PDEs with lognormal diffusion coefficient*, in Sparse Grids and Applications – Stuttgart 2014, Springer-Verlag, 2016.
- [33] F. NOBILE, R. TEMPONE, AND C. WEBSTER, *An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM Journal on Numerical Analysis, 46 (2008), pp. 2411–2442.
- [34] ———, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM Journal on Numerical Analysis, 46 (2008), pp. 2309–2345.
- [35] C. SCHILLINGS AND C. SCHWAB, *Sparse, adaptive Smolyak quadratures for Bayesian inverse problems*, Inverse Problems, 29 (2013). doi:10.1088/0266-5611/29/6/065011.
- [36] C. SCHWAB AND C. GITTELSON, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numerica, 20 (2011), pp. 291–467.
- [37] G. SZEGŐ, *Orthogonal Polynomials*, American Mathematical Society, New York, fourth ed., 1975.
- [38] D. XIU AND J. HESTHAVEN, *High-order collocation methods differential equations with random inputs*, SIAM Journal on Scientific Computing, 37 (2005), pp. 1118–1139.